

Processo Seletivo

Vaga: Estágio

Área: IBRE FGV

Etapas: 2

Maio de 2023

Instruções:

1. Salve a base de dados “*base_ipea.csv*” que você recebeu por e-mail, junto com este PDF, em uma pasta de sua preferência.
2. Escolha, entre os softwares R ou Python, aquele com o qual você tem mais familiaridade.
3. Leia atentamente o texto da próxima página e responda aos questionamentos.
4. Encaminhe suas respostas (tanto os resultados quanto os comandos utilizados para gerá-los) para o e-mail taise.lyra@fgv.br até às 12h (horário de Brasília) do dia 11/05/2023.

Será desclassificado (a) o (a) candidato (a) que:

- Enviar somente os resultados, sem o código do R/Python elaborado para alcançá-los ou
- Enviar as respostas depois das 12h (horário de Brasília) do dia 11/05/2023.

IMPORTANTE:

O tempo que o (a) candidato (a) levar para enviar as respostas também será avaliado e servirá como critério de desempate para classificação. Isto é, imagine que os candidatos A e B acertaram todas as questões, porém o candidato A levou 2 dias para enviar suas respostas e o candidato B levou 1. Neste caso, o candidato B ficará em primeiro lugar e o candidato A em segundo.

Observação:

Fique à vontade para realizar consultas a quaisquer materiais em busca de auxílio para responder aos questionamentos. Mas tenha em mente que, justamente por este motivo, nossa avaliação será extremamente rigorosa. Capriche! E boa sorte!

Durante os meses de maio e junho de 2013, o IPEA realizou a pesquisa intitulada de Tolerância social à violência contra a mulher”. A pesquisa foi feita por amostragem, ou seja, foi selecionada uma parte da população do Brasil. A amostra é composta por 3810 indivíduos de ambos os sexos. E abrange os municípios metropolitanos e não-metropolitanos das cinco regiões (Norte, Nordeste, Centro-Oeste, Sudeste e Sul). Na pesquisa, frases foram lidas para os entrevistados, que em seguida deveriam dizer se: concordavam total ou parcialmente, discordavam total ou parcialmente ou nem concordavam nem discordavam (neutralidade).

- Defina o seu diretório de trabalho para o local onde se encontra a base de dados.
- Importe a base de dados *base_ipea.csv* para o software escolhido.
- A base é composta por quantas linhas e colunas?
- Calcule a frequência de pessoas em cada região do Brasil.
- Qual é a região mais frequente (moda)?
- Qual é a idade da pessoa mais nova nessa amostra? E da mais velha?
- Calcule a média, a mediana e a moda para a variável idade. A partir disso, o que você pode dizer sobre a distribuição dessa variável (assimétrica positiva, assimétrica negativa ou simétrica)?
- Classifique as idades de acordo com as faixas etárias a seguir. (Crie uma nova coluna no data frame para essa classificação). A amostra é composta de mais Jovens, Adultos ou Idosos?

Idade (x)	Faixa Etária
$x \leq 29$	Jovens
$30 \leq x \leq 59$	Adultos
$x \geq 60$	Idosos

- Calcule a média, a mediana, o primeiro quartil, o terceiro quartil e os valores máximo e mínimo para a variável “renda total de todos os moradores, parentes e agregados no último mês”. Comente os resultados.
- Interprete o primeiro e o terceiro quartis encontrados no item anterior.
- Crie uma função que calcule o coeficiente de variação.
- Calcule o coeficiente de variação para a variável idade e renda. Compare os dois coeficientes de variação.
- Calcule o desvio-padrão para a renda de acordo com cada região do Brasil. Qual é a região que possui um comportamento mais homogêneo em relação à renda?